DOCUMENT RESUME

ED 110 480 TM 004 753

AUTHOR Evers, John W.
TITLE An Investigation of Goal-Free and Goal-Based Evaluation Strategies through Project Director and Evaluator Ratings.
PUB DATE [Apr 75]
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Academic Achievement; *Comparative Analysis; *Evaluation; *Evaluation Methods; Evaluation Needs; Objectives; Performance; Professional Personnel
IDENTIFIERS Goal Based Evaluation; *Goal Free Evaluation

ABSTRACT

Both goal-free and goal-based methodologies were operationalized. Subjects for the study were nominated by leaders in the evaluation field. Subjects were randomly selected and assigned to treatments. Sites were randomly assigned to either methodology. Reports generated from each were sent to project directors to be rated on criteria found in the evaluation literature. Results test two differences inferred from the goal-free literature and generate hypotheses for further research on evaluation. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED110480

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

John W. Evers

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

An Investigation of Goal-Free and Goal-Based Evaluation Strategies through Project Director and Evaluator Ratings

John W. Evers

The Evaluation Center
Western Michigan University
Kalamazoo, Michigan

Presented at the 1975 Annual Meeting
of the American Educational Research Association

April 1, 1975
Session 7.11

TM004753

INTRODUCTION[1]

Differing evaluation approaches have been conceptualized by Alkin (1969), Cronbach (1963), Hammond (1967), Provus (1971), Scriven (1967), Stufflebeam (1971), Stake (1967, 1974), and Tyler (1950). Each considers different variables, different roles and functions for the evaluator, and different situations where they are applicable. However, no matter which approach is considered, all have one linking element: a use of some standard for comparison. That is, evaluation strategies use something as a standard whether it is goals, judgments, norms, values, performance, needs, efficiency, or an ideal state of being.

It was beyond the scope of this study to investigate more than one type of evaluation standard: goals. Goals, or in more operational terminology called objectives, are one standard that evaluators can use to assess achievement. The PDK Committee (1971) referred to this use of goals and objectives as the congruence approach to evaluation. Glass (1969) called this approach the Tylerian Model. Stake (1974) called it preordinate evaluation. No matter which label the approach falls under, it is defined in terms of a comparison of achievement to some prespecified goal statement(s).

Both Stake (1967) and Scriven (1967) have focused on the difficulties associated with using prespecified goals as a standard to judge achievement. Stake reported that assessing the congruence between achievement, (or outcomes) and goals, (or intentions) was only one

---

[1]A more complete discussion can be found in John W. Evers, "A Comparative Analysis of Goal-Free and Goal-Based Evaluation Strategies Through Project Director and Evaluator Ratings." Unpublished Dissertation, Western Michigan University, 1975.

part of the evaluation process. For example, Stake explained that the evaluator must also search for side effects and incidental gains rather than narrowly report only goal achievement. Scriven stated his position more strongly. He said that evaluation proper must include, as an equal partner with the measuring of performance against goals, procedures for the evaluation of the goals (p.52). That is, if the goals were not worth achieving then it is uninteresting to see how well they are achieved. Scriven explained that it was more important to ask How good is the course? rather than Did the course achieve its goals? Scriven elaborated by saying that espoused goals of the project director were often not the implicit goals of his project. Moreover, it is not always the case that this kind of error should be corrected in favor of the espoused goals by revising the project or in favor of the implicit goals by revising the espoused goals (p. 54). Finally as if a premonition, Scriven said succinctly that the evaluator should see what the project does, and not bother with the question of whether it had good intentions (p. 60).

Goals are considered necessary for management and planning, but are unnecessary for evaluation, as Scriven (1972) later reported. Evaluation was defined as the assessment of merit. Merit was independent of intention. For the evaluator to be aware of the espoused intentions of the project developed a perceptual set that more often than not biased judgments of the project's real achievements. By ignoring espoused intentions (stated goals), Scriven said that the evaluator would have a greater chance of assessing the merit of the real effects of the project. Merit of the real effects of the project was to be assessed by a comparison to the demonstrated needs of the target population. This approach was called goal-free evaluation (GFE) by Scriven because merit was

determined independent (free) from the intended goals. As Scriven (1973) later explained the basis of a goal-free approach was that of an impartial, neutral referee. He explained that the goal-free approach does not penalize the project staff for failing to reach overambitious goals. It gives them credit for doing something worthwhile in getting halfway to those goals. It does not restrict them to credit for their announced contracts since educators often do more good in unexpected directions than in intended ones. The GFE approach preserves their chances in those directions. (p. 323)

As could be expected the goal-free approach caused considerable discussion throughout the evaluation profession. For instance, Stufflebeam (1972, p. 5) commented on the overall merit of the GFE approach by saying that the strategy is potentially useful, but far from operational and replicable. Because of its promise Stufflebeam believed that Scriven and others should further develop, test, and report the effects of GFE whatever they turn out to be. Concurrently Popham (1972, p. 7) reviewed the possibilities of GFE. Although the strategy was alluringly portrayed by Scriven, Popham reported he would have to wait until GFE was tried in real evaluation settings to see its effects.

Scriven (1974) explained that the National Science Foundation was going to fund an experiment between two evaluation strategies to better understand their effects. The goal-free methodology was one approach to be used in the study. The other was as Scriven called it, goal-based evaluation (GBE). However, this experiment did not take place because of diffuse projects requiring large teams, much interaction between teams, "leakage" to a goal-free team was too possibie, and possible dilution of method effects was too great. Scriven went on to say in conclusion that

> It will take only a few such experiments [between GFE and GBE] (some with our trainees doing it instead of ourselves, to better isolate some of our own individual differences) to give us a good picture of GFE. I [Scriven] think its (GFE) value will be demonstrated if it sometimes picks up something significant at a cost which makes the discovery well worthwhile. (p. 47)

The NSF "experiment" did not take place because conditions were not conducive to a fair test between GFE and GBE. As Stufflebeam noted, the goal-free approach had potential usefulness and that it should be further developed, tested, and its effects reported. Therefore, one purpose of the study being reported was to investigate the effects of goal-free and goal-based evaluation methodologies through project director and evaluator ratings of the evaluation process.

OVERVIEW OF THE STUDY

In order to investigate effects of the two methodologies, naturalistic testing conditions were used. Within an evaluation contract at the Evaluation Center at Western Michigan University, there were several projects funded at twelve private, four-year colleges throughout the Midwest. Each project needed to be reviewed in order to assess its current activities. Projects were diverse although centered around one program funding area. Rather than introduce a bias by choosing specific evaluators for a specific methodology (as was proposed in the NSF experiment), a more equitable and rigorous test was to use several evaluators selected from a larger population. Since a fair test would necessitate competent individuals, a nationally-recognized group of twenty-five evaluators was asked to recommend individuals.

These individuals were sent a letter in the Spring of 1974 asking for at least one or two recommendations given these criteria:

1. Currently practicing evaluation as either a graduate student, or practitioner in the field.

2. Able to commit approximately six days to the task. This included a one-day orientation and training session at the Evaluation Center.

3. Has a proven ability to operate as an independent, solo evaluator.

4. Writes with an insightful, unlabored style.

5. Located within a radius of approximately 600 miles from the Center in order to reduce travel costs.

Thirty-one individuals were recommended that could meet these criteria. Individual evaluators were then selected and assigned to either methodology in a random manner. For the sake of practicality, and at

the same time diversity, six evaluators were chosen. Three were assigned to each of the two evaluation methodologies.

It should be important to note that evaluators were not told about two groups being involved or that any comparative study of evaluation methodologies was occurring. This subject ignorance of the overall study taking place within a "routine consulting assignment" was maintained until a debriefing session after all evaluation activities were completed. It was agreed during debriefing that specific names would not be linked directly to evaluation reports, and that confidentiality would be provided through use of composites of information.

The following table provides a composite of the two groups:

Table 1

Composite Background of Evaluators by Groups

		Background Variables		
Group	n	Sex	Previous Experience[a]	Highest Degree
Goal-Free:	3	1 Female 2 Male	1 Extensive 2 Moderate 0 Little	1 Doctorate 2 Masters
Goal-Based:	3	1 Female 2 Male	1 Extensive 1 Moderate 1 Little	2 Doctorate 1 Masters

[a]Background variable of previous experience in evaluation was qualified by examining subject's self-reported evaluation experiences on a background data sheet and scoring as follows:

"Extensive experience" was equal to subject experiences in evaluation covering historically dated periods of time greater than three years.

"Moderate experience" was equal to subject experiences in evaluation covering historically dated periods of at least one to three years.

"Little experience" was equal to subject experiences in evaluation covering historically dated periods of less than one year.

It can be seen that the goal-free group had more previous evaluation experience, however, the goal-based group had more academic experience at the doctoral level. It was assumed that the groups were not different to the degree that either was biased, or that one had a clear advantage, over the other.

Both the goal-free and goal-based evaluators were trained in the appropriate methodology. Procedures were developed from existing literature. Directions for reporting findings, an overview of the evaluator's assignment, and a narrative discussion of the methodology was developed into a notebook format that the evaluator could use at the projects. Also, a checklist approach was built into each methodology so that some calibration across evaluators would be possible.

Careful review was given to Scriven's (1974, p. 7-32) "Checklist for the Evaluation of Products, Producers, and Proposals" for the purpose of using a checklist approach to GFE. Scriven acclaimed the validity and utility of the checklist approach. He reported that it had been reviewed and revised for improvement. He went on to add that the checklist approach was important and that users of the approach should accept responsibility of its improvement. With this in mind, the "Handbook for Traveling Observers, First Edition" was developed. In early development of this handbook, the Product Evaluation Checklist was adapted so that further illustration and clarification was added to its original content. Each checkpoint was expanded into a matrix that used "description,"

"evidence used," "judgment of adequacy," and"recommendations" as column headings. Row headings were specific sub-points under each of the checklist's thirteen checkpoints. Preceding the checklist, the handbook had narrative sections that explained the setting of the evaluation study and the conceptual overview of the evaluator's role.

Similarly, the goal-based approach was developed around a checklist so that parallelism would exist in the overall structures implemented. Rather than developing a new checklist approach to GBE, another version was adapted. In an earlier evaluation conducted by the Evaluation Center for the National Science Foundation, Stufflebeam's CIPP evaluation model had been developed into workable reporting format similar to the checklist approach developed by Scriven. Minor developments were added to that earlier version to allow similarity between the column headings used in the two checklists. The row headings and each checkpoint varied from the GFE approach.

It should be noted as a slight digression that Scriven's checklist approach was directly transferable to the GFE approach since this approach used a standard of "demonstrated needs" to judge overall achievement, a position directly related to the goal-free approach. Stufflebeam's adapted checklist approach was not as directly transferable to the goal-based approach. However, this particular study structured the evaluators to use Stufflebeam's evaluation model primarily as "preordinate" or goal-based evaluation.

To further understand the effects of the two methodologies, information was collected about the evaluation process itself. This process information was used to supplement ratings on outcomes of the

methodologies so that a more comprehensive discussion of effects would be possible. Both the evaluators and the project directors were asked to make a general rating of the quality of the particular evaluation process used. The evaluator's ratings took place as soon as possible after leaving the project site. Even though the ratings were general, omitting many specific questions, it was decided that a short form could gather useful information without invading the evaluator's professionalism or perceived competency. The project director's process ratings took place immediately after the evaluator had left the project, and before the project director received the written evaluation report. Collecting ratings prior to reading the report was done so that no biasing effect would be introduced into the process rating from ratings of the written report. Similar to the evaluator form of the process rating, the project director form was short and asked for a general assessment of the quality of the evaluation process used on-site.

IMPLICATIONS OF FINDINGS FOR EVALUATION THEORY AND PRACTICE

Scriven's main objection to the goal-based approach to evaluation was that awareness of intentions, and goals, caused a perceptual set that limited the evaluator's objectivity to look beyond initial intentions to side effects. Evaluators in the goal-free group were kept ignorant of goal-laden information for as long as possible. There was evidence that evaluators were not totally ignorant of goals throughout the on-site visitation. At the same time, there was evidence that the goal-free evaluators were not totally informed of goals. It is difficult to speculate the degree to which ignorance of goals was represented in the written evaluation reports. However, project directors reading the goal-free reports did rate them higher in objectivity than did project directors rating the goal-based reports. This finding supports Scriven's contention that goal-free evaluation can be a more objective approach than goal-based evaluation.

Both groups of evaluators reported projects visited were in early developmental stages with little activity beyond preliminary planning and structuring. Evaluators who had been given goal-laden information were in a position to review planning decisions and goals since that was the most that many projects had produced at that time. Goal-free evaluators had been denied goal-laden information prior to on-site visitation and had been informed through the handbook protocol and training session to make a sincere attempt to resist discussions of goals and intentions with the project staff. Since the goal-free group was somewhat resistant to goal-laden discussions and projects were early in development with

little or no activities beyond goal-structuring, the goal-free group could have been limited in the potential amount of information that was possible to review and report.

Scriven's (1973) discussion of the goal-free approach raised questions of overt client anxiety. Whereas, actual ratings in this study found project directors rated the two evaluation processes so that they were not significantly different from each other than might be expected by chance. Therefore, it might be suggested from this finding that if the evaluator makes no overt display of the goal-free differences in procedure to the client, or project director, the report could be judged as equal in overall quality to one generated from a goal-based approach. The goal-free process may be judged similar in quality to the goal-based process. The written evaluation report can be as credible, and more objective, than one produced by a goal-based approach.

Ratings on the process itself found that goal-free evaluators were significantly different from goal-based evaluators in self-ratings of the quality of the evaluation process. The two evaluator groups did not differ on ratings of rapport with the client. This might have been anticipated since these were recommended individuals and one area of competency should be skill in human relations. However, the goal-free group did rate themselves lower than did the goal-based group in efficiency of time spent on-site, confidence to implement the goal-free approach, and overall satisfaction with the on-site visitation.

In summary, it could be said that a type of goal-free evaluation can be operationalized as an alternative to the preordinate approach. In implementing a goal-free approach, it is a useful consideration to

have an evaluator function in a mediating position between the goal-free evaluator and project-related information in order to screen out goal-related information. The checklist approach to evaluation can be useful in allowing some minimal level of structure to exist within the methodology. However, forms of the checklist used in this study may be too rigorous, or inappropriate, in question areas for a project rather early in its activities. The overall quality of evaluation presented in a written report may not differ if highly competent individuals are recruited as evaluators. However, closer analysis on specific criteria can find the goal-free information as more objective than goal-based information. Findings support that the goal-free evaluators rate their own evaluation processes lower when using a goal-free approach. Therefore, the goal-free evaluator may be a greater source of anxiety than the evaluation client, or the evaluator using a preordinate evaluation approach. Scriven (1973) has said

> In evaluation, blind is beautiful. Remember that Justice herself is blind, and good medical research is double blind. The educational evaluator is severely handicapped by the impossibility of double blind conditions in most educational contexts. But he or she must still work very hard at keeping out prejudicial information. You can't do an evaluation without knowing what it is you're supposed to evaluate--the work--the treatment--but you do not need or want to know what it's supposed to do. (p. 323).

In essence this point still has important implications for evaluation. However, it should be qualified: Blind is beautiful in evaluation, but evaluator orientation and mobility can be anxiety-ridden by not having goal statements to use as perceptual set like eyes to help poke, prod, and sort through project achievement. There is a similarity between a man stricken blind and an evaluator assessing achievement independent of explicit intentions. Both are anxious until learning the necessary skills of independence.

References

Alkin, Marvin C. "Evaluation Theory Development" Evaluation Comment, 2, 1969.

Cronbach, Lee J. "Course Improvement through Evaluation" Teachers College Record, 64, May 1963.

Glass, Gene V. "The Growth of Evaluation Methodology: Research Paper no. 27." Boulder: Laboratory of Educational Research, University of Colorado, 1969 (mimeo).

Hammond, Robert L. "Evaluation at the Local Level" Tucson: EPIC Evaluation Center, 1967 (mimeo).

Popham, W. James "Results Rather than Rhetoric" Evaluation Comment, 3, 1972.

Provus, Malcolm L. Discrepancy Evaluation. Berkeley: McCutchan Publishing Company, 1971.

Scriven, Michael S. "The Methodology of Educational Evaluation" in Robert E. Stake (Ed.), Curriculum Evaluation, AERA Series on Evaluation, no. 1, Chicago: Rand McNally Company, 1967.

Scriven, Michael S. Evaluation Skills. AERA Training Cassette Tapes, Tape 6B. Washington, D.C.: AERA, 1971.

Scriven, Michael S. "Prose and Cons about Goal-Free Evaluation" Evaluation Comment, 3, 1972.

Scriven, Michael S. "Goal-Free Evaluation" in Ernest R. House (Ed.), School Evaluation: The Politics and Process. Berkeley: McCutchan Publishing Company, 1973.

Scriven, Michael S. "Evaluation Perspectives and Procedures" in W. James Popham (Ed.), Evaluation in Education: Current Applications, Berkeley: McCutchan Publishing Company, 1974.

Stake, Robert E. "The Countenance of Educational Evaluation" Teachers College Record, 68, 1967.

Stake, Robert E. "Responsive Evaluation." Paper presented at the 1974 Annual Meeting AERA, Chicago.

Stufflebeam, Daniel L., et al Educational Evaluation and Decision Making. Itasca, Illinois: F.E. Peacock Publishers, 1971.

Stufflebeam, Daniel L. "Should or Can Evaluation Be Goal-Free?" Evaluation Comment, 3, 1972.

Tyler, Ralph W. Basic Principles of Curriculum and Instruction. Chicago: University of Chicago Press, 1950.